

**GenCore version 5.1.6**  
**Copyright (c) 1993 - 2003 Compugen Ltd.**  
**OM protein - protein search, using sw model**  
**Run on:** June 27, 2003, 17:58:49 ; **Search time 29 Seconds**  
 (Without alignments)  
 106.576 Million cell updates/sec  
**Title:** US-09-300-612-1  
**Perfect score:** 84  
**Sequence:** 1 LKAMDPTPPLWIKTE 15  
**Scoring table:** BLOSUM62  
**Gapext 0.5**  
**Searched:** 671580 seqs, 206047115 residues  
**Total number of hits satisfying chosen parameters:** 671580  
**Minimum DB seq length: 0**  
**Maximum DB seq length: 200000000**  
**Post-processing:** Minimum Match 0%  
 Maximum Match 100%  
**Listing first 45 summaries**



RA	Houston D., Houston K.A., Howland T.J., Wei M.-H., Idegwam C., Kennison J.A., Ketchem K.A.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Lai J., Lai Z., Lai Z.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkhov G., Milash N.V., Mobarry C., Morris J., Mosherfson A.,
RA	Mount S.M., Moy Nelson M., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Pittman K.A., Nixon K., Nusskern D.R., Pacled J.M., Reese M.G.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Purj V., Shen H.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA	Shue B.C., Sliden-Klamoos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhang M., Zhang G., Zhao Q., Zhang L.,
RA	Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.,
RT	"The genome sequence of <i>Drosophila melanogaster</i> .";
RT	Science 287:2185-2195(2000).
DR	EMBL: AE003550; AAC050190; -.
DR	DRFlyBase: FBgn0011836; Tarif150.
DR	InterPro: IPR001930; Ala_peptidase.
DR	InterPro: IPR002052; N6_Mtase.
DR	PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
SQ	SEQUENCE 139210 MW: 964663361C47D105 CRC64;
Qy	Query Match Score 46; DB 5; Length 1219;
Db	Best Local Similarity 53.3%; Pred. No. 66;
	Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps
RESULT 5	
Q8SZR7	PRELIMINARY; PRT: 1221 AA.
ID	Q8SZR7
AC	Q8SZR7;
DT	01-JUN-2002 (TREMBL); 21, Created)
DT	01-JUN-2002 (TREMBL); 21, Last sequence update)
DT	01-JUN-2002 (TREMBL); 21, Last annotation update)
DN	LD43729p.
TA	TAF150.
OS	<i>Drosophila melanogaster</i> (Fruit fly).
OC	Pterygota; Metazoa; Arthropoda; Hexapoda; Insecta; Diptera; Brachycera; Muscomorpha;
OC	Ephydrioidea; Drosophilidae; Drosophila.
NCBI_TAXID	7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-BERKELEY;
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Frise E.,
RA	Champé M., Chavez C., Dorsett V., Dresenek D., Farfan D., Frise E.,
RA	George R., Gonzalez M.J., Guarín H., Kronmiller B., Li P., Liao G.,
RA	Miranda A., Mongall J.C., Munoz J., Paragás V., Park S., Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Celniker S.; Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AY070564; AAC080551; -.
SQ	SEQUENCE 139498 MW: C2DC066826B1AF6E CRC64;
Qy	Query Match Score 46; DB 5; Length 1221;
Db	Best Local Similarity 53.3%; Pred. No. 67;
	Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps
Qy	1 LKAMDPPTPLWIKTE 15
Db	618 LSAMDDSPVNWIRLD 634

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	DR	Putative permease ABC transporter protein.
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovidae;	GN	R00154 OR SMC02830.
OC	Bovidae; Bos.	OS	Rhizobium meliloti. (Sinorhizobium meliloti).
OX	[1]	OC	Bacteria: Proto bacteria; alpha subdivision; Rhizobiaceae group;
RP	SEQUENCE FROM N.A.	OC	Rhizobiaceae; Sinorhizobium.
RX	MEIDLINE=20219146; PubMed=10753906;	OX	NCBI_TaxID=382;
RA	Rattner A., Smallwood P.M., Nathans J.;	RN	[1]
RA	"Identification and characterization of all-trans-retinol dehydrogenase from photoreceptor outer segments: the visual cycle enzyme that reduces all-trans-retinal to all-trans-retinol.";	RP	SEQUENCE FROM N.A.
RL	J. Biol. Chem. 275:11034-11043 (2000).	RC	STRAIN=1021;
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.	RX	Medline=21396507; PubMed=11481430;
CC	(SDR) FAMILY.	RA	Capel D., Barloy-Hubler F., Gouzy J., Bothé G., Ampe F., Batut J., Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S., Godrie T., Goffeau A., Kahn D., Kiss E., Leloir V., Masuy D., Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U., Renard C., Thebaud P., Vandenhoul M., Weidner S., Gallibert F.;
DR	EMBL; AF229846; AAF6161.1; -.	RA	"Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021.";
DR	HSSP; P14061; IEDBS.	RT	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
DR	InterPro; IPR00198; ADH_short.	RL	DR
PRam	PF00106; adh_short; 1.	EMBL; AL591782; CAC41541.1; -.	
PRINTZ	PRINTZ; PRO0080; SDREFAMILY.	DR	InterPro; IPR000515; BPD_transp.
DR	PROSITE; PS00065; ADH_SHORT; UNKNOWN_1.	DR	Pfam; PF00528; BPD_transp.
KW	Oxidoreductase; Receptor.	KW	Complete proteome.
SEQUENCE	312 AA; 33956 MW; 36481039FF71874D CRC64;	SQ	SEQUENCE 377 AA; 41867 MW; D8E34131C7E14415 CRC64;
Query Match	53.6%; Score 45; DB 6; Length 312;	Query Match	52.4%; Score 44; DB 16; Length 377;
Best Local Similarity	57.1%;	Best Local Similarity	60.0%;
Matches	Pred. No. 24;	Matches	Pred. No. 41;
Matches	4; Mismatches 2; Indels 0;	Matches	2; Mismatches 2; Indels 0;
Qy	1 LKAMDPPPLWKT 14	Qy	6 PTPPLWIKTE 15
	: :  :  :		: :
Db	268 LKAMDPPGSLLVVT 281	Db	125 PTPFWLMS 134
RESULT 9		RESULT 11	
Q9NV61	PRELIMINARY; PRT; 125 AA.	Q9UZG1	PRELIMINARY; PRT; 695 AA.
ID	Q9NV61;	ID	Q9UZG1;
AC		AC	
DT	01-OCT-2000 (TREMBLrel. 15, Created)	DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	OVARC00060 protein.	DE	Hypothetical protein PAB1590.
OS	Homo sapiens (Human).	GN	PAB1590.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OS	Pyrococcus abyssi.
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX	NCBI_TaxID=9606;	OC	Pyrococcus.
RN	[1]	OX	NCBI_TaxID=29295;
RP	SEQUENCE FROM CARCINOMA;	RN	[1]
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Negai K., Sugano S., Ishibashi T., Fujimori K., Tana H., Kimata M., Watanabe M., Hiraoaka S., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Matsuo Y., Kaneko K.;	RP	SEQUENCE FROM N.A.
RA	"NEDO human cDNA sequencing project.";	RC	STRAIN=ORSAY;
RA	Submitted (FEB-2000) to the ENSEMBL/GenBank/DDBJ databases.	RA	Heilig R.;
DR	EMBL; AK001769; BAA91896.1; -.	RT	"Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution";
DR	Submited (JUL-1999) to the ENSEMBL/GenBank/DDBJ databases.	RL	Submitted (JUL-1999) to the ENSEMBL/GenBank/DDBJ databases.
SQ	SEQUENCE 125 AA; 13504 MW; F2D324918E3A70B1 CRC64;	DR	EMBL; AJ248286; CAB50098.1; -.
Query Match	52.4%; Score 44; DB 4; Length 125;	SQ	Hypothetical protein; Complete Proteome. SEQUENCE 695 AA; 79999 MW; 371EUE40210F551 CRC64;
Best Local Similarity	46.2%;	Query Match	52.4%; Score 44; DB 17; Length 695;
Matches	Pred. No. 13;	Best Local Similarity	54.5%;
Matches	4; Mismatches 3; Indels 0;	Matches	Pred. No. 78;
Qy	2 KAMDPPTPLWIKT 14	Qy	3 AMDPTPPLWIK 13
	:  :  :  :  :		:  :  :  :  :
Db	4 KSMGPAPPWRS 16	Db	561 AND PFPWMSE 571
RESULT 10		RESULT 12	
Q92T32	PRELIMINARY; PRT; 377 AA.	Q15070	PRELIMINARY; PRT; 1327 AA.
ID	Q92T32	ID	Q15070;
AC	Q92T32	AC	Q15070;
DT	01-DEC-2001 (TREMBLrel. 19, Created)	DT	01-JAN-1998 (TREMBLrel. 05, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE	KIAA0364	protein.				
GN	KIAA0364					
OS	Homo sapiens (Human)					
OC	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Eutheria; Primates; Catarhini; Hominidae; Homo.					
NCBI_TAXID	9606;					
OX						
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=BRAIN;					
RX	Medline=97349904; PubMed=9205641;					
RA	Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,					
RA	Tanaka A., Kotani H., Nomura N., Ohara O.;					
RT	*Prediction of the coding sequences of unidentified human genes. VII.					
RT	The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";					
DNA	Res. 4:141-150(1997)					
RL	PLN 4:141-150(1997)					
DR	EMBL: AB002362; BAA20819; 1; -.					
DR	HSSP: HSSP026; INR.					
DR	InterPro: IPR03598; Ig_C2.					
DR	InterPro: IPR03600; Ig_Like.					
DR	InterPro: IPR03006; Ig_MHC.					
DR	SMART: SM00408; Ig_C1; 1.					
DR	SMART: SM00410; Ig_Like; 7.					
KW	Immunoglobulin domain; MW; 147772 MW;					
SO	SEQUENCE 1327 AA; 147772 MW; 736C689FEC94D2E1 CRC64;					
Query Match	52.4%; Score 44; DB 4; Length 1327;					
Best Local Similarity	63.6%; Pred. No. 1.5e+02;					
Matches	7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;					
Qy	4 MDPTPPWIKT 14					
Db	: 25 MDPTPPWIKS 35					
RESULT 13						
Q946HB	PRELIMINARY;	PRT; 278 AA.				
ID	Q946HB					
AC	Q946HB					
DT	01-DEC-2001 (TREMBLrel. 19, Created)					
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)					
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)					
DE	Alpha-expansin.					
GN	Oryza sativa (Rice).					
OS	Eukaryota; Viridiplantae; Streptophytta; Embryophytta;					
OC	Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae;					
OC	Ehrhartoidae; Oryzeae; Oryzae.					
NCBI_TAXID	4530;					
OX						
RN	SEQUENCE FROM N.A.					
RX	Medline=21521059; PubMed=11641069;					
RA	Lee Y., Choi D., Kende H.,					
RT	*Expansin: ever-expanding numbers and functions. ";					
RL	Curr. Opin. Plant Biol. 4:527-532(2001).					
DR	InterPro: IPR00882; Pollen_allergen.					
DR	Pfam: PF01357; Pollen_allergen; 1.					
DR	ProDom: PDO02179; Pollen_allergen; 1.					
DR	SEQUENCE 278 AA; 30046 MW;					
SO	D78E729F091766FO CRC64;					
Query Match	51.2%; Score 43; DB 10; Length 278;					
Best Local Similarity	66.7%; Pred. No. 44;					
Matches	6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;					
Qy	5 DETPPWIK 13					
Db	: : 44 DETPPWIK 52					
RESULT 14						
Q8VM74	PRELIMINARY;	PRT; 354 AA.				
ID	Q8VM74;					
AC	Q8VM74					
DT	01-MAR-2002 (TREMBLrel. 20, Created)					
DT	01-JUN-2002 (TREMBLrel. 20, Last sequence update)					
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DE	Putative LuxA.					
OS	Rhizobium sp. (strain BR816).					
OC	Proteobacteria; Rhizobiaceae; alpha subdivision; Rhizobiaceae group;					
OC	Rhizobiaceae; Rhizobium.					
OX	NCBI_TAXID=48291;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRATN-BR816;					
RA	Luyten E., Verreth C., Vanderleyden J.;					
RA	Conservation of short-chain alcohol dehydrogenase in rhizobial spp. ";					
RA	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRATN-BR816;					
RA	Luyten E., Swinnen B., Verreth C., Vlaasse K., Dombrecht B., Vanderleyden J.;					
RA	Functional and structural analysis of a P450 homolog gene identified in Rhizobium sp. BR816. ";					
RA	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.					
DR	U26451; AAL61564.1; Bac_luciferase.					
DR	InterPro: IPR002103; Bac_luciferase.					
DR	Pfam: PF00295; bac_luciferase_1.					
KW	Plasmid.					
SQ	SEQUENCE					
QY	39367 MW;					
DB	167 PPPPLWI 12 167 PPPPIWI 173					
RESULT 15						
Q944ST	PRELIMINARY;	PRT; 406 AA.				
ID	Q944ST;					
AC	Q944ST;					
DT	01-DEC-2001 (TREMBLrel. 19, Created)					
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)					
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DE	Arabidopsis thaliana (Mouse-ear cress).					
OS	Eukaryota; Viridiplantae; Streptophytta; Embryophytta; Spermatophytina; Magnoliophyta; eudicots; Rosidae; eurosids II; Brassicaceae; Arabidopsis.					
OC	Arabidopsis thaliana (Mouse-ear cress).					
OX	NCBI_TAXID=3702;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C., Goldsmith A.D., Hayashizaki Y., Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Karlin-Neumann G., Kawai J., Liang B., Jones T., Kamiya A., Miranda M., Narusawa M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C., Tominaga M., Yamada K., Yamamura Y., Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R., Shinohara T., Davis R.W., Theologis A., Ecker J.R.,					
RA	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.					
DR	AF42554; AAL1548.1;					
DR	InterPro: IPR001356; Homeobox.					
DR	InterPro: IPR002913; START.					
DR	Pfam: PF001852; START.					
DR	ProDom: PDO0010; Homeobox; 1.					
DR	PROSITE; PS00077; Homeobox; 1.					
DR	PROSITE; PS00071; Homeobox; 1.					
DR	PROSITE; PS05071; Homeobox; 1.					
DR	PROSITE; PS05074; UNKNOWN_1.					

SQ	SEQUENCE	406 AA:	45986 MW:	045B92F725BD785C CRC64;
Query	Match	51.2%	Score 43;	DB 10; Length 406;
Best Local Similarity	46.7%;	Pred. No. 65;		
Matches	7; Conservative	3; Mismatches	5; Indels	0; Gaps 0;
Qy	1 LKAMDPPPLWIKTE 15			
	1 : ;           :			
Db	2225 LRLLQTNPPLWIKTD 239			

Search completed: June 27, 2003, 18:02:00  
Job time : 31 secs